Error Estimates for an Ocean General Circulation Model from Altimeter and Acoustic Tomography Data

DIMITRIS MENEMENLIS*

Department of Earth, Atmospheric, and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts

MICHAEL CHECHELNITSKY

Massachusetts Institute of Technology–Woods Hole Oceanographic Institute Joint Program, Cambridge, Massachusetts

(Manuscript received 26 March 1998, in final form 4 January 1999)

ABSTRACT

An offline approach is proposed for the estimation of model and data error covariance matrices whereby covariance matrices of model data residuals are ''matched'' to their theoretical expectations using familiar least-squares methods. This covariance matching approach is both a powerful diagnostic tool for addressing theoretical questions and an efficient estimator for real data assimilation studies.

Provided that model and data errors are independent, that error propagation is approximately linear, and that an observability condition is met, it is in theory possible to fully resolve covariance matrices for both model and data errors. In practice, however, due to large uncertainties in sample estimates of covariance matrices, the number of statistically significant parameters that can be estimated is two to three orders of magnitude smaller than the total number of independent observations.

The covariance matching approach is applied in the North Pacific (5°–60°N, 132°–252°E) to TOPEX/Poseidon sea level anomaly data, acoustic tomography data from the Acoustic Thermometry of Ocean Climate Project, and a GCM. A reduced state linear model that describes large-scale internal (baroclinic) error dynamics is constructed. Twin experiments suggest that altimetric data are ill suited to estimating the statistics of the vertical GCM error structure, but that such estimates can in theory be obtained using acoustic data.

The particular GCM integration exhibits a warming trend relative to TOPEX/Poseidon data of order 1 cm yr^{-1} corresponding to a peak warming of up to 0.2°C yr^{-1} in the acoustic data at depths ranging from 50 to 200 m. At the annual cycle, GCM and TOPEX/Poseidon sea level anomaly are in phase, but GCM amplitude is 2 cm smaller, with the error confined above 200-m depth. After removal of trends and annual cycles, the low-frequency/wavenumber (periods >2 months, wavelengths >16°) TOPEX/Poseidon sea level anomaly is order 6 cm^2. The GCM explains about 40% of that variance. By covariance matching, it is estimated that 60% of the GCM–TOPEX/Poseidon residual variance is consistent with the reduced state linear model.

1. Introduction

This paper concerns the determination of error statistics for oceanic general circulation models (GCMs) and data. Error estimates are needed to understand model and data deficiencies and are prerequisite to data assimilation studies (Bennett 1992; Wunsch 1996). While a lot of effort already goes into providing error estimates for oceanic data (e.g., Stammer and Wunsch 1994), little is known, quantitatively, about the skill of GCMs. And yet, misspecification of GCM errors can have disastrous consequences on data assimilation results (Dee 1995). The major difficulty has been a lack of global oceanic datasets of sufficient quality and duration to characterize the error statistics. With the advent of the TOPEX/Poseidon altimeter, which provides stringent tests for GCM errors at the surface (Fu and Smith 1996; Stammer et al. 1996), and of other large-scale ocean observation systems (e.g., the TOGA–TAO array; Hayes et al. 1991), and with the proliferation of oceanic data assimilation studies (Malanotte-Rizzoli 1996), it becomes urgent to establish a quantitive framework in which to examine GCM errors.

The questions addressed here are the following. Given a particular observation system, what components of a GCM's error structure can be determined? Is it possible to discriminate between data and model errors? Can the altimetric data resolve the internal GCM error structure? How much data is required? Finally, what is the ac-

* Current affiliation: Jet Propulsion Laboratory, Pasadena, California.

Corresponding author address: Dimitris Menemenlis, Jet Propulsion Laboratory, Mail Stop 300/323, 4800 Oak Dr., Pasadena, CA 91109.
E-mail: dimitri@pacific.jpl.nasa.gov

curacy of the resulting errors estimates? We ask the above questions in the context of a study wherein four years of TOPEX/Poseidon data, the first full year of Acoustic Thermometry of Ocean Climate (ATOC) data, and the Marshall et al. (1997a,b) GCM are used to estimate the large-scale (>1000 km), time-varying circulation and heat budget of the North Pacific (The ATOC Consortium 1998).

The problem of quantifying model errors using data is addressed with adaptive Kalman filters by the engineering community (for recent surveys see Isaksson 1988; Moghaddamjoo and Kirlin 1993). Some of these adaptive filter methods have been applied to meteorological (e.g., Dee et al. 1985) and, more recently, to oceanographic (Blanchet et al. 1997) data assimilation studies. No single method, however, is applicable to all situations, each method being a compromise between computational cost, convergence speed, and simplifying assumptions.

Following the suggestion of Blanchet et al. (1997), we started by testing the empirical algorithm of Myers and Tapley (1976) and a maximum-likelihood estimator inspired by Dee (1995). We concluded that neither method was well suited to our particular problem, the former because of slow convergence and the latter because of high computational cost (Chechelnitsky 1999). We adapt instead the offline approach proposed by Fu et al. (1993) (appendix A) and estimate system and measurement error covariance matrices by matching sample covariance matrices of GCM data residuals to their theoretical expectations. Our algorithm extends the approach of Fu et al. (1993) in the following ways: 1) we relax the assumption of independence between model simulation errors and the true state, 2) we use Green's functions to obtain a solution, 3) we exploit time-lagged correlations in the data, and 4) we provide uncertainty bounds for the estimates.

The proposed covariance matching approach is similar to methods described by Shellenbarger (1966) and Belanger (1974) but we use GCM data residuals directly rather than the innovation sequence (i.e., residuals between data and successive Kalman filter estimates). Innovation sequence approaches have been preferred by the engineering community because they are more readily amenable to online applications and to the tracking of slowly varying statistics in small-dimensioned systems. When first guess error statistics are accurate, the innovations will be less correlated (in time) than GCM data residuals and, therefore, the available information will collapse into a small number of lag covariance matrices.

For the large-dimensioned systems of interest to oceanographic studies, however, it is preferable to work with GCM data residuals directly for the following reasons. First, sample covariances can be computed offline thus avoiding the computational burden associated with repeated integrations of the Kalman filter. Second, model and data error covariance matrices are linearly related to those of GCM data residuals. By way of contrast, the innovation sequence variants of the algorithm require linearization about some first guess error statistics and, therefore, convergence is not guaranteed (Moghaddamjoo and Kirlin 1993). Finally, GCM data residuals contain information about absolute error variances while the innovation sequence can be used only to determine the relative ratio of model and data error variances. Although relative error ratios suffice for time stepping the Kalman filter, absolute error variances are required for obtaining a posteriori error statistics.

The remainder of this article is organized as follows. A statistical description of the problem and of the covariance matching algorithm appear in section 2. In section 3, the circulation and measurement models are introduced, and covariance matching is tested in a series of twin experiments. In section 4, fields diagnosed from the GCM are compared to TOPEX/Poseidon and ATOC data and the residuals are used to estimate trends, annual cycles, and error covariance matrices. Conclusions are set forward in section 5.

2. Statistical modeling

Let $\mathbf{p}(t)$ represent GCM simulation errors,

$$\mathbf{p}(t) = \mathbf{x}_{\mathrm{GCM}}(t) - \mathbf{x}_{\mathrm{ocean}}(t), \tag{1}$$

that is, the difference at time t between $\mathbf{x}_{\mathrm{GCM}}(t)$, the prognostic variables of a GCM, and $\mathbf{x}_{\mathrm{ocean}}(t)$, the true state of the ocean sampled in a manner consistent with $\mathbf{x}_{\mathrm{GCM}}$. We model the dynamical evolution of the errors as

$$\mathbf{p}(t+1) = \mathbf{A}(t)\mathbf{p}(t) + \mathbf{B}(t)\mathbf{q}(t), \tag{2}$$

where $\mathbf{A}(t)$ is the state transition matrix and $\mathbf{q}(t)$ are system errors, that is, errors in initial and boundary conditions, indeterminate GCM parameters, and other model errors; $\mathbf{q}(t)$ is projected onto the GCM grid by matrix $\mathbf{B}(t)$. The difference between GCM predictions and oceanographic observations, $\mathbf{y}_{\mathrm{ocean}}(t)$, can be expressed as a noisy linear (or linearized) combination of $\mathbf{p}(t)$,

$$\mathbf{y}(t) = \mathbf{H}(t)\mathbf{x}_{\mathrm{GCM}}(t) - \mathbf{y}_{\mathrm{ocean}}(t), \tag{3}$$

$$= \mathbf{H}(t)\mathbf{p}(t) - \mathbf{r}(t), \tag{4}$$

where $\mathbf{H}(t)$ is the measurement matrix and $\mathbf{r}(t)$ are data errors. In addition to instrument noise, $\mathbf{r}(t)$ includes representation errors (Fukumori 1999; Cohn 1997), that is, real oceanic signal not represented by the GCM, for example, tides and scales smaller than those resolved by the model.

Vectors $\mathbf{q}(t)$ and $\mathbf{r}(t)$ are taken to be random variables and are described by their means, $\langle \mathbf{q}(t) \rangle$ and $\langle \mathbf{r}(t) \rangle$, and by their covariance matrixes, $\mathbf{Q}(t) \equiv \mathrm{cov}\mathbf{q}(t)$ and $\mathbf{R}(t) \equiv \mathrm{cov}\mathbf{r}(t)$, where the covariance operator is defined in the usual way, $\mathrm{cov}\mathbf{q} \equiv \langle (\mathbf{q} - \langle \mathbf{q} \rangle)(\mathbf{q} - \langle \mathbf{q} \rangle)' \rangle$, $\langle \cdot \rangle$ is the expectation operator, and prime indicates the trans-

TABLE 1. Summary of notation.

Symbol	Definition	Equations
$\tilde{\cdot}$; $\bar{\cdot}$; $\langle\cdot\rangle$, $\lvert\cdot\rvert$	Estimate, sample mean, expectation, norm	5, 15
$'$, cov, $(:)$	Transpose, covariance, column operator	5, 8, 11
p, σ^2, ρ	Degrees of freedom, variance, correlation coefficient	Appendix C
q, r, s, t, u, v	Time and time lag indices	1, 12, C1
K, L	Number of parameters: for $\mathbf{Q}$, for $\mathbf{R}$	6, 7
M, N	Vector length: $\mathbf{y}$, $\mathbf{p}$	Appendix B
S, T	Maximum delay, number of time steps	13, 15
$\mathbf{d}$	Vector of elements from $\mathbf{Y}$ and $\mathbf{D}_s$	13, 17
$\mathbf{r}$, $\mathbf{q}$, $\mathbf{p}$	Errors: data, system, GCM	1, 2, 4
$\mathbf{x}_{\text{ocean}}$, $\mathbf{x}_{\text{GCM}}$, $\mathbf{y}_{\text{ocean}}$, $\mathbf{y}$	Truth, GCM, data, residual	1, 3, 4
$\mathbf{A}$, $\mathbf{B}$, $\mathbf{H}$	Linear models	2, 4
$\mathbf{D}_s$, $\mathbf{Y}_s$	Lag-difference and lag covariance matrices	12, 14
$\mathbf{I}$, $\mathbf{0}$	Identity matrix; zero matrix	5, 12
$\mathbf{G}_{\mathbf{Y},k}$, $\mathbf{G}_{\mathbf{D}_s,k}$	Green's function	10, 13
$\mathbf{P}$, $\mathbf{Q}$, $\mathbf{R}$, $\mathbf{Y}$	cov$\mathbf{p}$, cov$\mathbf{q}$, cov$\mathbf{r}$, cov$\mathbf{y}$	6, 7, 8, 9
$\mathbf{P}_\alpha$, $\mathbf{R}_\alpha$, $\mathbf{R}_\epsilon$	cov$\tilde{\boldsymbol{\alpha}}$, cov$\boldsymbol{\alpha}$, cov$\boldsymbol{\epsilon}$	18
$\boldsymbol{\alpha}$, α_k, $\boldsymbol{\epsilon}$	Parameters for $\mathbf{Q}$ and $\mathbf{R}$, sample error	6, 7, 13, 17, 24
$\mathcal{G}$	Green's function kernel matrix	13, 17

pose. This is a complete statistical description of the errors if the random vectors $\mathbf{q}(t)$ and $\mathbf{r}(t)$ have multivariate normal distribution (e.g., Mardia et al. 1979), that is, if the errors can be modeled as resulting from a set of stationary Gaussian processes. If the errors are non-Gaussian, the mean and covariance remain useful, though incomplete, descriptors. Our objective is to use measurements $\mathbf{y}(t)$ to estimate $\langle\mathbf{q}(t)\rangle$, $\langle\mathbf{r}(t)\rangle$, $\mathbf{Q}(t)$, and $\mathbf{R}(t)$ (see Table 1 for a summary of the notation).

a. The basic algorithm

We start by considering the case where $\mathbf{A}$, $\mathbf{B}$, $\mathbf{H}$, $\mathbf{Q}$, and $\mathbf{R}$ are time-independent; $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{H}$ are known; $\mathbf{A}$ is stable, that is, all its eigenvalues are contained within the unit circle; $\mathbf{B}$ is identity; and vectors $\mathbf{q}(t)$ and $\mathbf{r}(t)$ have zero mean and are independent of $\mathbf{p}(t)$,

$$\langle\mathbf{q}(t)\rangle = \mathbf{0}, \qquad \langle\mathbf{r}(t)\rangle = \mathbf{0},$$
$$\langle\mathbf{p}(t)\mathbf{q}(t)'\rangle = \mathbf{0}, \qquad \langle\mathbf{p}(t)\mathbf{r}(t)'\rangle = \mathbf{0}. \tag{5}$$

[The assumption of independence between $\mathbf{q}(t)$ and $\mathbf{p}(t)$ is less restrictive than that used by Fu et al. (1993), who assumed the model simulation error to be independent of the true state, $\langle\mathbf{p}(t)\mathbf{x}_{\text{ocean}}(t)'\rangle = \mathbf{0}$.] For stable $\mathbf{A}$, (2), (4), and (5) imply that $\langle\mathbf{y}(t)\rangle = \langle\mathbf{p}(t)\rangle = \mathbf{0}$; (2) and (5) imply that $\langle\mathbf{q}(t_1)\mathbf{q}(t_2)'\rangle = \mathbf{0}$ for $t_1 \neq t_2$. Finally we parameterize $\mathbf{Q}$ and $\mathbf{R}$ as

$$\mathbf{Q} = \sum_{k=1}^{K} \alpha_k \mathbf{Q}_k, \tag{6}$$

$$\mathbf{R} = \sum_{k=1}^{L} \alpha_{K+k} \mathbf{R}_k. \tag{7}$$

Equations (6) and (7) anticipate that only a small number of parameters can be resolved with a reasonable degree of statistical significance. The exact forms of $\mathbf{Q}_k$ and $\mathbf{R}_k$ are problem dependent. Ideally, they should approximate the leading spatial patterns, or eigenvectors, of the errors. In practice they are chosen based on physical intuition and using Occam's razor, that is, a search for the simplest, physically plausible, and statistically consistent error model.

Multiplying (2) by its transpose and taking expectations produces the steady state Lyapunov equation (e.g., Anderson and Moore 1979, p. 62),

$$\mathbf{P} \equiv \text{cov}\mathbf{p} = \mathbf{A}\mathbf{P}\mathbf{A}' + \mathbf{Q}, \tag{8}$$

which relates the covariance of the GCM error to that of the system error. For stable $\mathbf{A}$, the Lyapunov equation is readily solved for $\mathbf{P}$ using any of a number of analytical or numerical schemes (Gajić and Qureshi 1995). Similarly, multiplying (4) by its transpose and taking expectations yields

$$\mathbf{Y} \equiv \text{cov}\mathbf{y} = \mathbf{H}\mathbf{P}\mathbf{H}' + \mathbf{R}. \tag{9}$$

From (8) and (9) it follows that each element of $\mathbf{Y}$ is linearly related to the elements of $\mathbf{Q}$ and $\mathbf{R}$, and hence to parameters α_k in (6), (7). An elegant way to solve this system of equations is through the use of Green's functions, $\mathbf{G}_{\mathbf{Y},k}$, here defined as the response of measurement covariance matrix, $\mathbf{Y}$, to unit perturbations of $\mathbf{Q}_k$ or $\mathbf{R}_k$, that is,

$$\mathbf{G}_{\mathbf{Y},k} = \mathbf{H}\mathbf{P}_k\mathbf{H}', \qquad \mathbf{G}_{\mathbf{Y},K+k} = \mathbf{R}_k, \tag{10}$$

where $\mathbf{P}_k$ is related to $\mathbf{Q}_k$ by the Lyapunov equation (8). Rewriting $\mathbf{Y}$ and $\mathbf{G}_{\mathbf{Y},k}$ as column vectors yields a set of linear equations,

$$\mathbf{Y}(:) = [\mathbf{G}_{\mathbf{Y},1}(:) \cdots \mathbf{G}_{\mathbf{Y},K+L}(:)] \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_{K+L} \end{bmatrix}, \tag{11}$$

which can be solved for parameters α_k using any of several discrete linear inverse methods (e.g., Menke 1989; Wunsch 1996). To reduce computational cost, the

column operator $(\vdots)$ in (11) can also represent an appropriate subsampling of matrices $\mathbf{Y}$ and $\mathbf{G}_{\mathbf{Y},k}$, for example, their diagonal elements. For any given definition of operator $(\vdots)$ and set of matrices $\mathbf{A}$ and $\mathbf{H}$, linear inverse theory provides powerful tools for understanding which, and how well, combinations of parameters α_k in (11) can be determined.

This completes a basic description of the estimation algorithm. We next consider a series of algorithmic refinements and the effects of relaxing some of the simplifying assumptions. One issue is whether $\mathbf{R}$ and $\mathbf{Q}$ can be estimated simultaneously (e.g., Groutage et al. 1987; Maybeck 1979), that is, whether an arbitrary set of parameters α_k in (11) can be resolved independently. In section 2b and appendix B we demonstrate that, under a very general set of conditions, $\mathbf{R}$ and $\mathbf{Q}$ can be resolved by making use of time-lagged correlations in the data. A more serious limitation is that $\mathbf{Y}$ is estimated as the sample covariance of $\mathbf{y}(t)$: the consequences of sampling uncertainty are discussed in section 2c and appendix C. The algorithm is illustrated with a small numerical example in section 2d. Systematic and time-correlated errors are considered in sections 2e and 2f, respectively. Section 2g deals with time-dependent models. Finally, section 2h discusses statistical consistency tests.

b. Using lag-difference covariance matrices

The covariance matrix $\mathbf{Y}$ does not describe temporal correlations in the data. It is therefore reasonable to expect that estimates of $\mathbf{Q}$ and $\mathbf{R}$ might be improved by making use of lag-difference covariance matrices. From a recursive application of (2) and from the definitions of $\mathbf{Q}$, $\mathbf{R}$, and $\mathbf{P}$, the covariance matrix of the lag-s difference is

$$\begin{aligned}\mathbf{D}_s &\equiv \mathrm{cov}[\mathbf{y}(t+s) - \mathbf{y}(t)] \\ &= 2\mathbf{HPH}' - \mathbf{HA}^s\mathbf{PH}' - \mathbf{HPA}^{s\prime}\mathbf{H}' + 2\mathbf{R} \\ &= \mathbf{H}(\mathbf{A}^s - \mathbf{I})\mathbf{P}(\mathbf{A}^s - \mathbf{I})'\mathbf{H}' \\ &\quad + \sum_{k=1}^{s} \mathbf{HA}^{s-k}\mathbf{QA}^{(s-k)\prime}\mathbf{H}' + 2\mathbf{R},\end{aligned} \tag{12}$$

where it is assumed that $\langle \mathbf{r}(t_1)\mathbf{r}(t_2)'\rangle = \mathbf{0}$ for $t_1 \neq t_2$. As before, $\mathbf{Y}$ and several lag-s-difference covariance matrices can be combined in an equation of type $\mathbf{d} = \mathcal{G}\boldsymbol{\alpha}$; that is,

$$\begin{bmatrix} \mathbf{Y}(\vdots) \\ \mathbf{D}_1(\vdots) \\ \vdots \\ \mathbf{D}_S(\vdots) \end{bmatrix} = \begin{bmatrix} \mathbf{G}_{\mathbf{Y},1}(\vdots) & \cdots & \mathbf{G}_{\mathbf{Y},K+L}(\vdots) \\ \mathbf{G}_{\mathbf{D}_1,1}(\vdots) & \cdots & \mathbf{G}_{\mathbf{D}_1,K+L}(\vdots) \\ & \vdots & \\ \mathbf{G}_{\mathbf{D}_S,1}(\vdots) & \cdots & \mathbf{G}_{\mathbf{D}_S,K+L}(\vdots) \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_{K+L} \end{bmatrix}. \tag{13}$$

Here, $\mathbf{G}_{\mathbf{D},k}$ represents the Green's function associated with data covariance matrix $\mathbf{D}$ and parameter α_k. Since $\mathbf{D}_r - \mathbf{D}_s$ is independent of $\mathbf{R}$ for any $r \neq s$, it is possible to resolve a particular $\mathbf{Q}_k$ independently of $\mathbf{R}$ (see appendix B), provided $\mathbf{Q}_k$ is observable in the sense that $\mathbf{HA}^s\mathbf{Q}_k\mathbf{A}^{s\prime}\mathbf{H}' \neq \mathbf{0}$ for some $s \geq 1$.

An equation of type (13) can also be written in terms of the lag covariance,

$$\mathbf{Y}_s \equiv \langle \mathbf{y}(t+s)\mathbf{y}(t)'\rangle = \mathbf{HA}^s\mathbf{PH}'. \tag{14}$$

Whether it is preferable to use lag rather than lag-difference covariance matrices, that is, $\mathbf{Y}_s$ rather than $\mathbf{D}_s$, is addressed in the next section.

c. Finite number of measurements

The discussion so far has assumed that covariance matrices $\mathbf{Y}_s$ and $\mathbf{D}_s$ are exact. In practice, a finite number of measurements is available and we work with sample estimates $\tilde{\mathbf{Y}}_s$ and $\tilde{\mathbf{D}}_s$: the sample covariance of $\mathbf{y}(t)$ is

$$\tilde{\mathbf{Y}} \equiv \frac{1}{T}\sum_{t=1}^{T} [\mathbf{y}(t) - \overline{\mathbf{y}}][\mathbf{y}(t) - \overline{\mathbf{y}}]', \tag{15}$$

where T is the total number of time steps and

$$\overline{\mathbf{y}} \equiv \frac{1}{T}\sum_{t=1}^{T} \mathbf{y}(t) \tag{16}$$

is the sample mean.

The first algorithmic modification required concerns the computation of Green's functions. If T spans less than about 20 e-folding periods for each observable normal mode of linear system $\mathbf{p}(t+1) = \mathbf{Ap}(t)$, the steady-state limit given by the solution to the Lyapunov equation (8) will be inaccurate. A Monte Carlo approach can instead be used to estimate $\mathbf{P}_k$ by driving linear model (2) with random system noise generated using covariance $\mathbf{Q}_k$; $\mathbf{P}_k$ is estimated by averaging over a large number of independent simulations, each with finite time span T.

A second modification is required to accomodate uncertainty in the sample covariance matrices. This is achieved by appending an error term, vector $\boldsymbol{\epsilon}$, to (13):

$$\mathbf{d} = \mathcal{G}\boldsymbol{\alpha} + \boldsymbol{\epsilon}. \tag{17}$$

To solve (17), it is appropriate to use variance-minimizing methods because the probability distribution of sample covariance matrices is approximately normal (appendix C). For example, parameter vector $\boldsymbol{\alpha}$ in (17) can be determined by minimizing the weighted least-squares cost function,

$$J(\boldsymbol{\alpha}) = \boldsymbol{\epsilon}'\mathbf{R}_\epsilon^{-1}\boldsymbol{\epsilon} + (\boldsymbol{\alpha} - \boldsymbol{\alpha}_0)'\mathbf{R}_\alpha^{-1}(\boldsymbol{\alpha} - \boldsymbol{\alpha}_0), \tag{18}$$

where $\boldsymbol{\alpha}_0$, $\mathbf{R}_\alpha$, and $\mathbf{R}_\epsilon$ represent prior knowledge for $\langle\boldsymbol{\alpha}\rangle$, cov$\boldsymbol{\alpha}$, and cov$\boldsymbol{\epsilon}$, respectively.

The uncertainty variance of a sample covariance is $O[\sigma_1^2\sigma_2^2(1+\rho^2)/p]$, where σ_1^2 and σ_2^2 denote variances for the two random variables being compared, ρ is the correlation coefficient, and p is the number of degrees of freedom, that is, the number of independent measurements (appendix C). It follows that for a given sample size, the smaller the variances, the more accurately sample covariances can be determined.

For example, in the twin experiments of section 3c, statistically significant error estimates are possible using lag-difference covariance matrices, $\mathbf{D}_s$, but not with lag covariance matrices, $\mathbf{Y}_s$. In those experiments the errors propagate slowly relative to the duration of a time step, that is, the state transition matrix is approximately identity, so that, for small values of lag s, (12) simplifies to $\mathbf{D}_s \approx s\mathbf{HQH}' + 2\mathbf{R}$. The sample uncertainty of $\tilde{\mathbf{D}}_s$ therefore scales with the diagonal elements of ($s\mathbf{HQH}' + 2\mathbf{R}$). By comparison, the uncertainty of $\tilde{\mathbf{Y}}_s$ scales with the diagonal elements of ($\mathbf{HPH}' + \mathbf{R}$), which are much larger. As a rule of thumb, it is preferable to work with $\tilde{\mathbf{D}}_s$ when $\mathbf{A} \approx \mathbf{I}$ and $|\mathbf{R}| \ll |\mathbf{HPH}'|$.

d. Numerical example

The covariance matching recipe is next illustrated using a numerical example. Consider the system of equations (2), (4) with

$$\mathbf{A} = \begin{bmatrix} 0.8 & 0.2 \\ -0.1 & 0.9 \end{bmatrix}, \qquad \mathbf{B} = \mathbf{I}, \qquad \mathbf{H} = [1 \quad 1]. \tag{19}$$

The system and measurement error covariance matrices are parameterized

$$\mathbf{Q} = \alpha_1 \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + \alpha_2 \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} + \alpha_3 \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix},$$
$$\mathbf{R} = \alpha_4. \tag{20}$$

From the steady-state Lyapunov equation (8) we obtain the covariance matrices

$$\mathbf{P}_1 = \begin{bmatrix} 2.5 & -0.4 \\ -0.4 & 0.5 \end{bmatrix}, \qquad \mathbf{P}_2 = \begin{bmatrix} 1.9 & 1.7 \\ 1.7 & 3.7 \end{bmatrix},$$
$$\mathbf{P}_3 = \begin{bmatrix} 5.9 & 3.2 \\ 3.2 & 2.5 \end{bmatrix}, \tag{21}$$

corresponding to unit perturbations of parameters α_1, α_2, and α_3, respectively, in (20). Computing the Green's functions associated with $\mathbf{Y}$ and $\mathbf{D}_s$ results in the following system of equations:

$$\begin{bmatrix} \mathbf{Y} \\ \mathbf{D}_1 \\ \mathbf{D}_2 \\ \mathbf{D}_3 \end{bmatrix} = \begin{bmatrix} 2.2 & 9.1 & 14.9 & 1 \\ 1.3 & 1.1 & 4.4 & 2 \\ 2.3 & 2.6 & 8.6 & 2 \\ 3.2 & 4.4 & 12.6 & 2 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \end{bmatrix}. \tag{22}$$

The kernel matrix in (22) has rank 3 (singular values $[24.6 \quad 3.7 \quad 0.9 \quad 0.0]'$), which indicates that only three independent combinations of parameters α_k can be resolved. It turns out that the addition of $\mathbf{D}_3$, or of higher lag covariance matrices, does not contribute new information. Rules regarding the total number of resolvable parameters are set forward in appendix B.

Simulated data were generated for $\boldsymbol{\alpha} = [1 \quad 1 \quad 0 \quad 1]'$, $T = 500$. We seek to estimate $\boldsymbol{\alpha}$ using the simulated data and the recipe of section 2b. From inverse theory, only projections onto singular vectors of $\mathcal{G}$ corresponding to nonzero singular values can be determined (e.g., Wunsch 1996, p. 147). The full solution is

$$\tilde{\boldsymbol{\alpha}} = [0.08 \quad 0.25 \quad 0.47 \quad 1.00]' + \lambda[-0.88 \quad -0.34 \quad 0.34 \quad 0.00]', \tag{23}$$

where λ is an arbitrary constant multiplying null space contributions; λ cannot be determined without additional information. To set λ we assume that there is a priori knowledge that the system error covariance matrix is diagonal, that is, $\alpha_3 = 0$. This assumption requires that $\lambda = -1.4$ and hence that $\tilde{\boldsymbol{\alpha}} = [1.3 \quad 0.7 \quad 0 \quad 1.0]'$.

Next we seek to estimate the solution uncertainty, $\mathbf{P}_\alpha \equiv \mathrm{cov}\tilde{\boldsymbol{\alpha}}$. Formally $\mathbf{P}_\alpha$ is a function of a priori covariance matrices $\mathbf{R}_\epsilon$ and $\mathbf{R}_\alpha$ in (18). Here $\mathbf{R}_\alpha$ is the a priori covariance of parameter vector $\boldsymbol{\alpha}$ and the only a priori knowledge assumed is that $\alpha_3 = 0$. Matrix $\mathbf{R}_\epsilon$ describes the sample uncertainty of $\tilde{\mathbf{Y}}$ and $\tilde{\mathbf{D}}_s$. An estimate of $\mathbf{R}_\epsilon$, consistent with the available data, can be obtained using the expressions derived in appendix C:

$$\mathbf{R}_\epsilon \approx \begin{bmatrix} 1.8 & 0.1 & 0.2 & 0.5 \\ 0.1 & 0.1 & 0.1 & 0.1 \\ 0.2 & 0.1 & 0.2 & 0.3 \\ 0.5 & 0.1 & 0.3 & 0.6 \end{bmatrix}. \tag{24}$$

The solution uncertainty matrix is $\mathbf{P}_\alpha = (\mathcal{G}'\mathbf{R}_\epsilon^{-1}\mathcal{G})^{-1}$. Therefore $\tilde{\boldsymbol{\alpha}} = [1.3 \pm 0.4 \quad 0.7 \pm 0.2 \quad 0 \quad 1.0 \pm 0.2]'$, consistent with the parameter vector $\boldsymbol{\alpha} = [1 \quad 1 \quad 0 \quad 1]'$ used to generate the simulated data. (Unless otherwise specified, uncertainty is reported using one standard deviation.)

From a set of numerical experiments, like that above, we conclude that the covariance matching method gives consistent and statistically significant estimates, provided the total number of available measurements is much greater than the number of parameters α_k, that is, $MT \gg K + L$, where M is the length of the measurement vector, T is the number of time steps, and $K + L$ is the total number of parameters in (6), (7). The requirement for a large number of observations per parameter is a direct consequence of the large uncertainty of sample covariance matrices.

What happens if instead of assuming $\alpha_3 = 0$, which is the condition used to generate the simulated data, it is instead assumed that $\alpha_1 = 0$? This assumption implies that $\lambda = 0.09$ in (23) and leads to a second solution $\tilde{\boldsymbol{\alpha}} = [0 \quad 0.2 \pm 0.3 \quad 0.5 \pm 0.2 \quad 1.0 \pm 0.2]'$. From the data alone there is no way to decide whether this solution is better or worse than the previous one. In fact, there exist a large number of consistent solutions depending on particular choices of λ and of other a priori assumptions. For this particular example, a second independent measurement at every time step would permit $\mathbf{Q}$ to be determined uniquely. But for real oceanographic

problems there is rarely, if ever, sufficient data to fully determine $\mathbf{Q}$, and one must therefore rely on physical intuition to choose suitable models for the errors.

(MATLAB script files and functions that implement this example, and which can be customized for different applications, are available via anonymous FTP to gulf.mit.edu, IP Address 18.83.0.149, from directory pub/dimitri/GCMerror.)

e. Systematic errors

Systematic errors, or biases, refer to the quantities $\langle \mathbf{r}(t) \rangle$ and $\langle \mathbf{q}(t) \rangle$. These errors are important because, even if very small, they can accumulate over long numerical integrations and degrade the predictive skill of a model. A first scenario is that of a stable, time-independent system, as before, but with $\langle \mathbf{r} \rangle \neq \mathbf{0}$, $\langle \mathbf{q} \rangle \neq 0$. Notice that the estimators that have been developed for $\mathbf{R}$, $\mathbf{Q}$, and $\mathbf{P}$ are not, to first order, affected by the presence of measurement and model biases because the sample mean is subtracted from the data in (15) and because the biases cancel out when computing lagged data differences.

Model bias correction in the context of atmospheric data assimilation was recently discussed by Evensen et al. (1998) and by Dee and da Silva (1998). They described online algorithms suitable for sequential estimation approaches. Offline algorithms, whereby biases are removed prior to data assimilation, are also available (e.g., Fukumori 1999) and are discussed below for completeness. From (2) and (4) it follows that

$$\langle \mathbf{y} \rangle = \mathbf{H}(\mathbf{I} - \mathbf{A})^{-1} \langle \mathbf{q} \rangle + \langle \mathbf{r} \rangle, \tag{25}$$

that is $\langle \mathbf{y} \rangle$ is linearly related to the biases, $\langle \mathbf{q} \rangle$ and $\langle \mathbf{r} \rangle$. The sample mean, $\overline{\mathbf{y}}$ in (16), is an unbiased estimator of $\langle \mathbf{y} \rangle$ with uncertainty (Anderson 1971, section 8.2),

$$\mathrm{cov}\overline{\mathbf{y}} = \sum_{r=1-T}^{T-1} \left(\frac{T - |r|}{T^2} \right) \langle \mathbf{y}(t + r) \mathbf{y}(t)' \rangle, \tag{26}$$

which, using (2), (4), and (8), reduces to

$$\mathrm{cov}\overline{\mathbf{y}} = \frac{\mathbf{Y}}{T} + \sum_{r=1}^{T-1} \left(\frac{T - r}{T^2} \right) (\mathbf{H}\mathbf{A}^r \mathbf{P} \mathbf{H}' + \mathbf{H}\mathbf{P}\mathbf{A}^{r\prime} \mathbf{H}'), \tag{27}$$

where $\mathbf{Y}$ is the data covariance matrix (9) and $\mathbf{P}$ is the GCM error covariance matrix (8); estimates for both matrices having been obtained earlier. Without additional information or assumptions, it is not possible to discriminate between system bias $\langle \mathbf{q} \rangle$ and data bias $\langle \mathbf{r} \rangle$.

A second scenario, that of a gradual change, or trend, in the system error, is discussed in section 2g, which deals with time-dependent models.

f. Time-correlated errors

So far we have assumed that measurement and system errors are uncorrelated in time, that is $\langle \mathbf{r}(t_1) \mathbf{r}(t_2)' \rangle = \mathbf{0}$, $\langle \mathbf{q}(t_1) \mathbf{q}(t_2)' \rangle = \mathbf{0}$ for $t_1 \neq t_2$. The former condition is required to evaluate lag-difference covariance matrices (12), but it is not required to evaluate the data covariance matrix $\mathbf{Y}$, which can therefore be used as before. The latter condition is implicit in (5) and presents a more difficult modeling challenge. Correlated system and data errors need to be removed from the model-data residuals in order to avoid biasing estimates of second-order statistics.

Under the Kalman filter formalism this situation is usually addressed by appending additional parameters to the state vector and jointly estimating time-correlated and uncorrelated errors. These parameters can also be estimated offline. Consider, for example, the specific case of an annual cycle in the system and/or in the measurement errors, a situation that is of direct practical relevance to oceanographic applications. Taking the Fourier transform of (2) and (4), it follows that each frequency component of $\mathbf{y}(t)$ is linearly related to the same frequency component in $\mathbf{q}(t)$ and $\mathbf{r}(t)$,

$$\mathbf{y}_a = \mathbf{H}\mathbf{p}_a + \mathbf{r}_a, \tag{28}$$

$$\mathbf{p}_a \exp(i\omega/12) = \mathbf{A}\mathbf{p}_a + \mathbf{q}_a, \tag{29}$$

where the subscript a indicates the complex annual cycle amplitude, that is, $\mathbf{y}_a = \mathbf{a} \exp(i\phi)$, $\mathbf{a}$ is the amplitude, ϕ is the phase, $\omega = 2\pi$/year, and we have assumed a time step of 1 month in (29). It is straightforward to remove correlated signals at the annual period from model-data residuals (e.g., section 4). But without additional information, it is not possible to partition the annual cycle error between system and data errors.

g. Time-dependent models

We consider two types of time dependence. The first type is ''known'' time dependencies in the linear models, $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{H}$, and also possibly in the measurement error covariance matrix, $\mathbf{R}$. These are readily accommodated by using a Monte Carlo approach to compute the Green's functions, $\mathbf{G}_{\mathbf{D},k}$. An example of this approach, with a time-varying $\mathbf{H}(t)$, is the treatment of acoustic time series of differing lengths in section 4.

The second type of time dependence is fluctuations of the ''unknown'' model parameters, $\alpha_k(t)$ in (6), (7). In principle, this situation can be addressed through piecewise estimates of $\alpha_k(t)$ for periods that are short relative to the timescales of α_k. A better approach is to parameterize the time dependency and to estimate these parameters using all the available data. An example is the detection of a trend, $\langle \partial \mathbf{q}/\partial t \rangle \neq \mathbf{0}$, in the system error. From (25), and assuming $\langle \partial \mathbf{r}/\partial t \rangle = \mathbf{0}$, the first difference of $\mathbf{q}(t)$ is related to the first difference of $\mathbf{y}(t)$ by

$$\langle \mathbf{y}(t + 1) - \mathbf{y}(t) \rangle = \mathbf{H}(\mathbf{I} - \mathbf{A})^{-1} \langle \mathbf{q}(t + 1) - \mathbf{q}(t) \rangle. \tag{30}$$

The expression $\langle \mathbf{y}(t + 1) - \mathbf{y}(t) \rangle$ can be approximated using least-squares (or other suitable estimators) and in turn used to estimate the quantity $\langle \partial \mathbf{q}/\partial t \rangle$ (e.g., section 4).

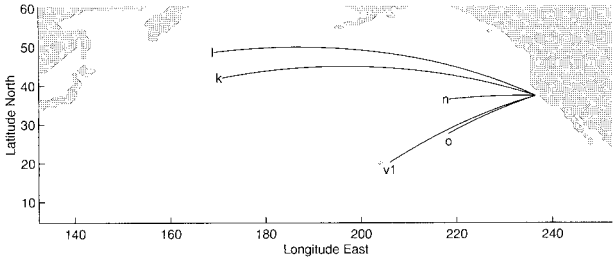

FIG. 1. Acoustic sections from the Acoustic Thermometry of Ocean Climate (ATOC) experiment used in the analysis are overlain on a map of the study area.

h. Tests of consistency

The final step of any estimation study is to test the resulting estimates for statistical consistency with all a priori assumptions. One possible test is the comparison of estimation residual, $\boldsymbol{\epsilon}$ in (17), (18), to its expected a posteriori covariance,

$$\text{cov}\tilde{\boldsymbol{\epsilon}} = \mathbf{R}_\epsilon[\mathbf{R}_\epsilon(\mathcal{G}\mathbf{R}_\alpha\mathcal{G}' + \mathbf{R}_\epsilon)^{-1}]'. \qquad (31)$$

In addition, when $\mathbf{Q}$ and $\mathbf{R}$ are used in conjunction with a Kalman filter, whiteness tests can be applied to the innovation vectors (Daley 1992).

The description of the algorithm is now complete. In the remainder of this article we illustrate the application of this algorithm, first with twin experiments (section 3) and then with real data (section 4), by estimating the large-scale (>1000 km) baroclinic errors in a particular implementation and linearization of a GCM.

3. Twin experiments

a. Circulation and measurement models

The circulation and measurement models, described below, are common to both the twin and the real experiments. The GCM is that of Marshall et al. (1997a,b) integrated in a global configuration wih realistic topography and driven by surface wind and buoyancy fields obtained from twice-daily National Centers for Environmental Prediction (NCEP) meteorological analyses. Horizontal grid spacing is 1° and there are 20 vertical levels.

A linear, time-independent model for GCM errors in the North Pacific is constructed by systematically perturbing the GCM with large-scale temperature anomalies (Menemenlis and Wunsch 1997). The linear model is defined in a region bounded by 5°–60°N and 132°–252°E (Fig. 1). It operates on a reduced state vector that has 8° sampling in the horizontal, four vertical temperature empirical orthogonal functions (EOFs, see Fig. 2), and a time step of 1 month. In this representation, sea surface pressure errors in the GCM caused by barotropic or salinity effects, or by scales not resolved by the reduced state vector, become part of the measurement error $\mathbf{r}(t)$, and are described by covariance matrix $\mathbf{R}$. The state vector dimension is reduced from 5×10^6 in the GCM to 512 in the linear model. Away from coastal regions, this reduced-state linear model describes the large-scale temperature perturbation response of the GCM with considerable skill for periods up to two years. Similar types of state reduction and linearization are commonly used for propagating the error covariance matrix in data assimilation studies (Fukumori and Malanotte-Rizzoli 1995; Cane et al. 1996).

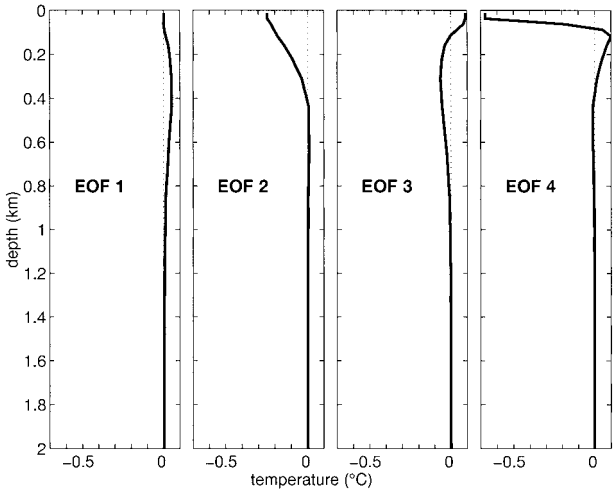

FIG. 2. Temperature EOFs used to represent the vertical GCM error structure. EOFs were computed from the difference between GCM output and measured temperature profiles in the study area (Menemenlis et al. 1997).

The acoustic tomography data from ATOC are first inverted to produce equivalent range-averaged oceanographic temperature perturbations along each section (The ATOC Consortium 1998). Data–GCM discrepancy is then projected onto the four vertical EOFs and the monthly sampling of the reduced state vector described above. Therefore, the measurement matrix for acoustic tomography data consists of a range-average for each vertical EOF and for each section. Acoustic data from five sections (Fig. 1) are used for a total of 20 data points (projections onto the four vertical EOFs for each section), once per month.

The measurement matrix $\mathbf{H}$ appropriate for altimetry consists of a weighted sum of the four vertical EOFs at each horizontal location of the reduced state grid. The weights are chosen to represent steric sea level anomalies corresponding to unit perturbations of each vertical EOF at each location.

b. Generation of simulated data

Before applying covariance matching to real data, we test the algorithm in a series of twin experiments using simulated data with known statistical properties. We parameterize $\mathbf{Q}$ in (6) as a diagonal matrix with four parameters, $\alpha_1, \ldots, \alpha_4$, each representing system error variance associated with each of the four vertical EOFs, that is, we assume that the system error is horizontally

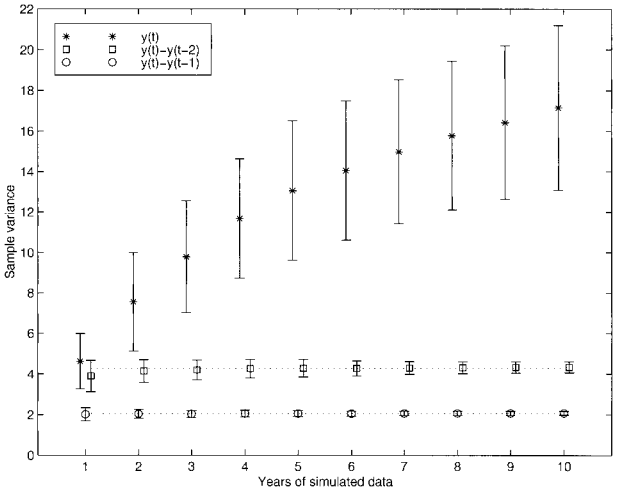

FIG. 3. Mean diagonal values of sample covariance matrices $\tilde{\mathbf{Y}}$, $\tilde{\mathbf{D}}_1$, and $\tilde{\mathbf{D}}_2$, as a function of years of simulated data. Error bars represent the associated standard uncertainty. Dotted lines are the steady-state values. (The steady-state value associated with $\tilde{\mathbf{Y}}$, 30, is not shown. Because the leading eigenvalue of the reduced-state dynamical model corresponds to an *e*-folding timescale of 19 years, a few hundred years of data are needed for sample covariance $\tilde{\mathbf{Y}}$ to reach a steady value). Lag-1 estimates have the smallest relative uncertainty, suggesting that matching the lag-1 data covariance matrix will provide the most accurate estimates of system and measurement error.

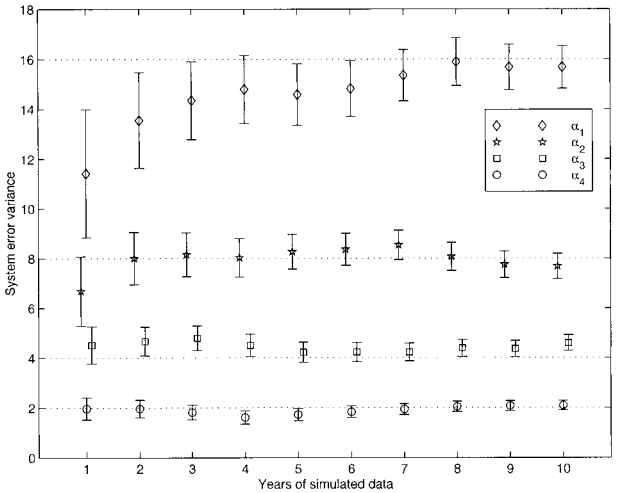

FIG. 4. Estimates of system error variance based on the lag-1 difference sample covariance, $\tilde{\mathbf{D}}_1$, for simulated acoustic tomography data. Dotted lines indicate the values of $\alpha_1, \ldots, \alpha_4$ used to generate the test data. The error bars represent the standard error of the estimates. The figure demonstrates the increasing accuracy of the algorithm with increasing number of measurements.

homogeneous and white. The measurement error covariance, $\mathbf{R}$ in (7), is also modeled as a diagonal matrix with two parameters, α_5 and α_6, corresponding to the measurement error variance associated with acoustic tomography and altimeter data, respectively. The test data are generated using the reduced state linear model and the acoustic and altimetric measurement models, and by driving Eqs. (2), (4) with white system and measurement noise characterized by parameters $\alpha_1, \ldots, \alpha_6$, as defined above.

c. Tests with pseudoacoustic data

The first set of twin experiments is carried out with noise-free, $\mathbf{R} = \mathbf{0}$, simulated acoustic tomography data. It is both impractical, because of computational cost, and unnecessary, because of information overlap, to match all available lag-difference data covariance matrices as in (13). An appropriate subset of data covariance matrices must be selected by trial and error and by reference to the guidelines of section 2c, that is, a preference for sample covariance matrices with small matrix norms and hence smaller relative uncertainties. The sample uncertainties of $\tilde{\mathbf{Y}}$, $\tilde{\mathbf{D}}_1$, and $\tilde{\mathbf{D}}_2$ are displayed in Fig. 3 as a function of number of years of simulated data. Note that $\tilde{\mathbf{D}}_1$ and $\tilde{\mathbf{D}}_2$ have smaller relative uncertainties than $\tilde{\mathbf{Y}}$, suggesting that matching $\tilde{\mathbf{D}}_1$ or $\tilde{\mathbf{D}}_2$ will produce better estimates of $\mathbf{Q}$ and $\mathbf{R}$ than matching $\tilde{\mathbf{Y}}$.

Figure 4 displays estimates of parameters $\alpha_1, \ldots, \alpha_4$, based on matching $\tilde{\mathbf{D}}_1$, as a function of number of years of simulated data. Error bars are obtained as in section 2d. Contrary to the empirical algorithm of Myers and Tapley (1976), which failed to converge for this twin experiment (see Chechelnitsky 1999), the present algorithm provides useful estimates of system error even with a single year of data.

The results of a series of tests based on 14 months of simulated data are summarized in Table 2 (at the time of this study 14 months of ATOC data were available). Each particular estimate is not expected to match the true variance of $\mathbf{Q}$ exactly, but over a large number of realizations the estimates are unbiased and their standard deviation matches the standard uncertainty reported on Table 2. For the dynamical and measurement models used here, the lag-1 difference sample covariance matrix, $\tilde{\mathbf{D}}_1$, provides the most accurate estimates, with mean standard uncertainty of 18% as compared to 38% for $\tilde{\mathbf{Y}}$. Matching only the diagonal elements of $\tilde{\mathbf{D}}_1$ leads to a standard uncertainty of 23% similar to that obtained by using the full lag-2 difference covariance matrix, $\tilde{\mathbf{D}}_2$.

Next we report on results from a series of experiments with noisy measurements, $\mathbf{R} = \mathbf{I}$ (Table 3). Measurement

TABLE 2. Estimates of system error covariance matrix $\mathbf{Q}$ based on 14 months of simulated acoustic tomography data. The measurements are assumed perfect, that is, $\mathbf{R} = \mathbf{0}$. The best estimates are obtained by matching the lag-1 difference covariance matrix, $\tilde{\mathbf{D}}_1$, with an average standard error of 18% as compared to 38% for $\tilde{\mathbf{Y}}$.

Estimate description	Parameters of covariance matrix $\mathbf{Q}$			
	α_1	α_2	α_3	α_4
Truth	16	8	4	2
From $\tilde{\mathbf{Y}}$	10.8 ± 5.7	9.2 ± 3.2	5.5 ± 1.6	2.1 ± 0.7
From $\tilde{\mathbf{D}}_1$	12.6 ± 2.7	8.0 ± 1.4	5.3 ± 0.8	1.9 ± 0.4
From diag$\tilde{\mathbf{D}}_1$	11.8 ± 3.4	8.8 ± 1.8	5.2 ± 0.9	1.7 ± 0.5
From $\tilde{\mathbf{D}}_2$	12.5 ± 3.5	5.2 ± 1.8	4.2 ± 0.9	1.6 ± 0.5

TABLE 3. Estimates of system and measurement error variance based on 14 months of simulated acoustic tomography data with $\mathbf{R} = \mathbf{I}$. The addition of measurement error increases the uncertainty of the estimates as compared to those of Table 2. Nevertheless, usable estimates of $\mathbf{Q}$, with a standard error of 38%, are possible by simultaneously matching the lag-1 and lag-2 difference covariance matrices.

Estimate description	Parameters of covariance matrix $\mathbf{Q}$ and $\mathbf{R}$				
	α_1	α_2	α_3	α_4	α_5
Truth	16	8	4	2	1
From $\tilde{\mathbf{Y}}$	10.4 ± 6.2	10.5 ± 3.6	5.8 ± 2.1	4.0 ± 1.3	0.6 ± 0.3
From $\tilde{\mathbf{D}}_1$	10.9 ± 4.5	12.7 ± 3.1	0.4 ± 2.4	−0.2 ± 1.6	1.3 ± 0.2
From $\tilde{\mathbf{D}}_2$	8.8 ± 4.7	11.3 ± 2.9	4.3 ± 2.1	2.9 ± 1.4	1.0 ± 0.2
From $\tilde{\mathbf{D}}_1$ and $\tilde{\mathbf{D}}_2$	8.9 ± 3.8	9.5 ± 2.4	6.0 ± 1.7	2.0 ± 1.1	1.0 ± 0.1

error degrades the estimates of $\mathbf{Q}$ considerably: the standard error for estimates obtained using $\tilde{\mathbf{D}}_1$ is 52%. The uncertainty can be reduced by using several lag-s difference covariance matrixes simultaneously: using $\tilde{\mathbf{D}}_1$ and $\tilde{\mathbf{D}}_2$ simultaneously reduces the estimation uncertainty to 38%.

In summary, the estimation uncertainty decreases with increasing years of available data and with increasing ratio $|\mathbf{Q}|/|\mathbf{R}|$. The simulation results indicate that 14 months of acoustic data are sufficient to produce usable estimates of $\mathbf{Q}$ and $\mathbf{R}$, provided the circulation and measurement models of 3a are valid and provided $|\mathbf{Q}| \gg |\mathbf{R}|$.

d. Tests with pseudoaltimeter data

A third set of twin experiments is conducted using simulated altimeter data. In theory, it is possible to separate baroclinic modes in the altimeter data by making use of their different temporal evolutions at the sea surface (e.g., Holland and Malanotte-Rizzoli 1989). The results presented below, however, suggest that even with perfect measurements, $\mathbf{R} = \mathbf{0}$, and with perfect knowledge of the dynamical and measurement models, $\mathbf{A}$ and $\mathbf{H}$, altimeter data on their own are ill-suited to the estimation of the vertical GCM error statistics. Figure 5 is an attempt to estimate the system error using up to 10 years of perfect altimeter data. At the conclusion of year 10, the standard uncertainty of the estimates remains too large for the estimates to be of any practical interest.

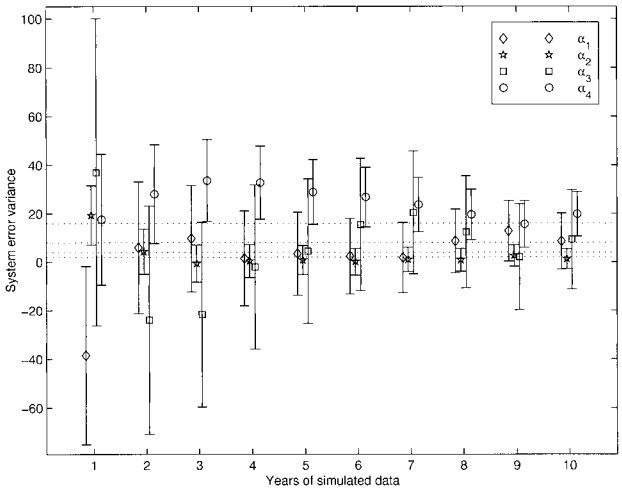

FIG. 5. Estimates of system error variance based on the diagonal elements of $\tilde{\mathbf{D}}_1$ for simulated altimeter data. Dotted lines indicate variances used to generate the data. Error bars represent the standard uncertainty of the estimates and they can be compared to those of Fig. 4, which was created using simulated acoustic data. The large error bars associated with the altimetric estimates suggest that altimeter data are ill-suited to the estimation of the vertical GCM error structure.

At the writing of this manuscript, 48 months of high quality TOPEX/Poseidon altimeter data were available. We therefore performed a further series of tests using 48 months of simulated altimeter data (see Table 4). Because of the large dimensions of the sample covariance matrices, only their diagonal elements have been matched. The first six rows of Table 4 correspond to estimates from matching $\mathbf{Y}$, and $\mathbf{D}_1$ through $\mathbf{D}_5$. The last row summarizes results from matching all six data covariance matrices simultaneously. The standard errors for this last case range from 35% to 235%. The situation is worse when measurement errors are included. We conclude that covariance matrices for the vertical GCM error structure cannot, in the present setup, be quantified from TOPEX/Poseidon data alone.

4. Experimental results

a. TOPEX/Poseidon data

The covariance matching approach is next applied to TOPEX/Poseidon altimeter data and to a particular integration of the Marshall et al. (1997a,b) GCM. Figure 6 compares measured sea level anomaly variance to that predicted by the GCM. Both the altimetric data and the GCM have been processed in a way consistent with the reduced state described in section 3a, that is, periods shorter than 2 months and length scales smaller than 16° have been low-pass filtered. In addition, annual cycles and trends have been removed at every location; these will be studied separately. Altimetric data and GCM output exhibit the same general patterns of enhanced variability near the Kuroshio, the Hawaiian Ridge, and in a band north of the equator. The GCM variability, however, is on average 30% less than that measured by the altimeter, and in some regions, notably

TABLE 4. Estimates of system error covariance matrix $\mathbf{Q}$ based on 48 months of perfect, $\mathbf{R} = \mathbf{0}$, simulated altimeter data. The last row of numbers are estimates obtained using the diagonal elements from all six data covariance matrices, $\tilde{\mathbf{Y}}$, $\tilde{\mathbf{D}}_1$, . . ., $\tilde{\mathbf{D}}_5$, simultaneously.

Estimate description	Parameters of covariance matrix $\mathbf{Q}$ and $\mathbf{R}$			
	α_1	α_2	α_3	α_4
Truth	16	8	4	2
From diag($\tilde{\mathbf{Y}}$)	30 ± 14	15 ± 8	−1 ± 20	−15 ± 9
From diag($\tilde{\mathbf{D}}_1$)	−1 ± 16	2 ± 7	−2 ± 30	32 ± 14
From diag($\tilde{\mathbf{D}}_2$)	30 ± 16	7 ± 7	−41 ± 28	10 ± 12
From diag($\tilde{\mathbf{D}}_3$)	35 ± 14	18 ± 8	−45 ± 28	−10 ± 11
From diag($\tilde{\mathbf{D}}_4$)	36 ± 13	22 ± 8	−36 ± 29	−19 ± 11
From diag($\tilde{\mathbf{D}}_5$)	38 ± 13	27 ± 9	−35 ± 30	−26 ± 10
From all the above	8 ± 6	10 ± 3	15 ± 9	3 ± 4

in the eastern tropical Pacific, the altimetric and GCM time series are uncorrelated. The variance of the GCM–TOPEX/Poseidon residual (Fig. 6c) is 60% that of TOPEX/Poseidon, indicating that the GCM explains 40% of the observed low-frequency/wavenumber variability. Our objective is to determine which fraction of the GCM–TOPEX/Poseidon residual can be attributed to system error, $\mathbf{HPH}'$ in (9), and which fraction results from measurement and representation errors, $\mathbf{R}$ in (9).

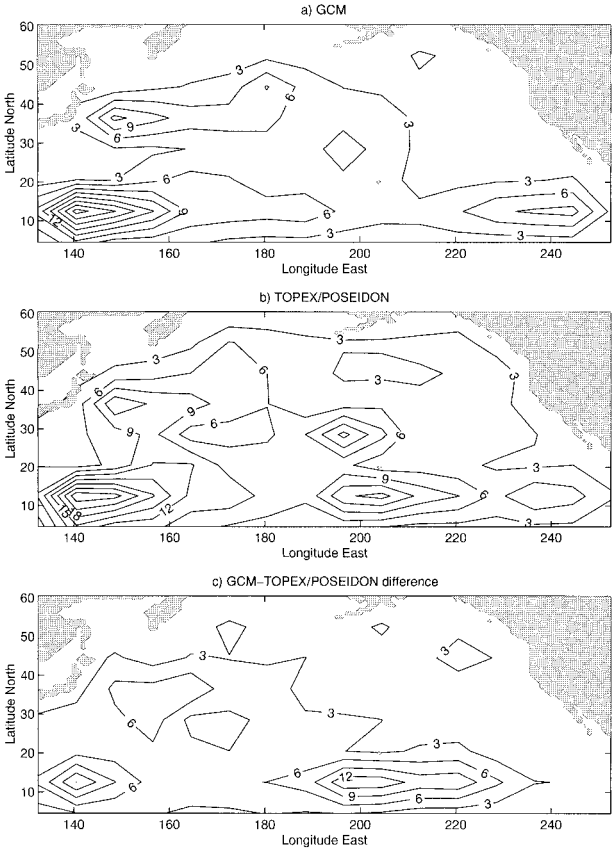

FIG. 6. North Pacific sea level anomaly variance for (a) GCM output, (b) TOPEX-Poseidon data, and (c) GCM-TOPEX/Poseidon residual, during the period 1 October 1992–31 May 1997. Annual cycles, trends, periods shorter than 2 months, and length scales smaller than 16° have been removed. Contour intervals are 3 cm^2.

The twin experiments conducted earlier indicate that it is not possible to determine covariance matrices for the vertical GCM error structure from four years of altimetric data. We therefore consider a number of statistical models for covariance matrices $\mathbf{Q}$ and $\mathbf{R}$ that assume equipartition of the variance between the four vertical EOFs. The first model is an attempt to estimate the full spatial structure of the error variance under the assumption that $\mathbf{Q}$ and $\mathbf{R}$ have zero off-diagonal elements. This model results in estimates that have no statistical significance; on average the standard uncertainty of the estimates is 15 times larger than the estimates themselves for the diagonal elements of $\mathbf{Q}$ and two times larger for the diagonal elements of $\mathbf{R}$.

To obtain statistically significant error estimates, it is necessary to reduce the number of parameters to be estimated. Therefore the second model considered is one of homogeneous and spatially uncorrelated system and measurement error, $\mathbf{Q} = \alpha_1\mathbf{I}$ and $\mathbf{R} = \alpha_2\mathbf{I}$, repectively. Matching this model to covariance matrices $\tilde{\mathbf{Y}}$, $\tilde{\mathbf{D}}_1$, $\tilde{\mathbf{D}}_2$, and $\tilde{\mathbf{D}}_3$ yields $\tilde{\alpha}_1 = 0.25 \pm 0.02$, $\tilde{\alpha}_2 = 1.00 \pm 0.03$. Standard uncertainties are computed using a set of 100 Monte Carlo experiments whereby covariance matching is applied to 100 sets of simulated data generated using normally distributed $\mathbf{q}(t)$ and $\mathbf{r}(t)$ with variance 0.25 and 1.00, respectively. Assuming the statistical model chosen to be the correct one, the standard deviation of the Monte Carlo estimates represents a lower bound for the standard uncertainty of the real estimates. These estimates imply that on average 70% of the GCM–TOPEX/Poseidon residual variance can be explained by system error; that is, the ratio of the diagonal elements of $\mathbf{HPH}'$ in (9) to those of $\mathbf{Y}$ is approximately 70% (see Fig. 7a).

The homogeneous model, however, does not account for some of the regions of enhanced variability in Fig. 6c. A third plausible model is $\mathbf{Q} = \alpha_1\mathbf{Q}_1$ and $\mathbf{R} = \alpha_2\mathbf{R}_1$, where $\mathbf{Q}_1$ and $\mathbf{R}_1$ are diagonal matrices with a spatially varying structure proportional to that of the GCM–TOPEX/Poseidon residual variance. Matching this model to the data yields $\tilde{\alpha}_1 = 0.047 \pm 0.006$, $\tilde{\alpha}_2 = 0.28 \pm 0.02$, which implies that 60% of the GCM–TOPEX/Poseidon residual variance is explained by system error (Fig. 7b). To within the sample and estimation uncer-

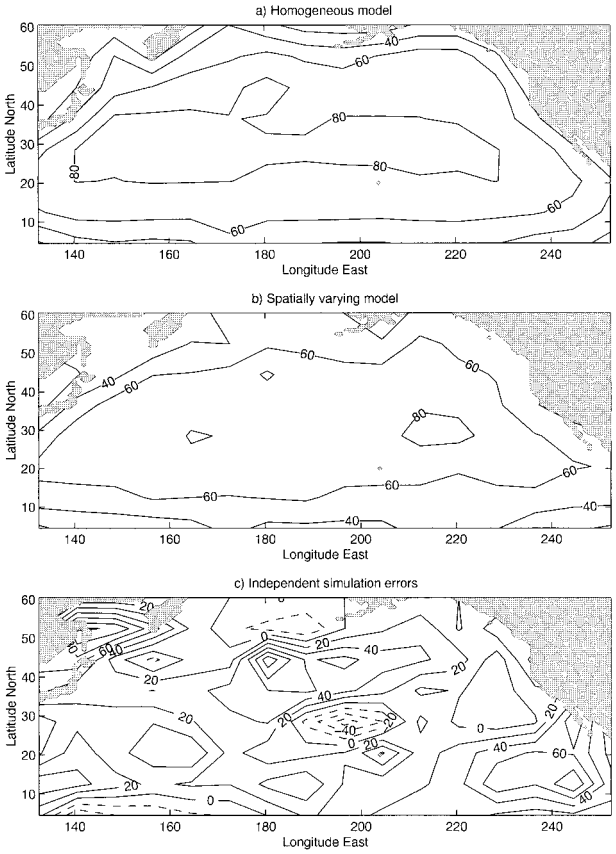

FIG. 7. Prior estimate for percent variance of GCM-TOPEX/Poseidon residual, which is explained by system error, that is, ratio of diagonal elements of $\mathbf{HPH}'$ in (9) to those of $\mathbf{Y}$. The estimates are obtained using the covariance matching method for (a) a homogeneous model for the errors, (b) a spatially varying model, and (c) the model proposed by Fu et al. (1993), which assumes that the GCM simulation errors are independent from the ocean state. Contour intervals are 20%. Spurious negative regions in (c) (dashed contours) result from the large uncertainty of the sample covariance matrices used in the analysis.

tainties, the prior variance predicted by this second model is consistent with the data.

A fourth model, that proposed by Fu et al. (1993), results from assuming that the ocean state is independent from the GCM simulation error, $\langle \mathbf{x}_{\text{ocean}}\mathbf{p}' \rangle = 0$ (see appendix A). When this assumption holds,

$$\mathbf{R} = \frac{1}{2}(\mathbf{Y} + \text{cov}\mathbf{y}_{\text{ocean}} - \mathbf{H}\,\text{cov}\mathbf{x}_{\text{GCM}}\mathbf{H}'), \quad (32)$$

$$\mathbf{HPH}' = \frac{1}{2}(\mathbf{Y} - \text{cov}\mathbf{y}_{\text{ocean}} + \mathbf{H}\,\text{cov}\mathbf{x}_{\text{GCM}}\mathbf{H}'). \quad (33)$$

On average, this third model predicts that 15% of the GCM–TOPEX/Poseidon residual variance is caused by system error (Fig. 7c). Although this relatively low value, compared to the earlier estimate of 60%, could point to a number of problems with the statistical model, the presumption is that to first-order condition $\langle \mathbf{x}_{\text{ocean}}\mathbf{p}' \rangle = \mathbf{0}$ is violated and hence that the 15% estimate is wrong.

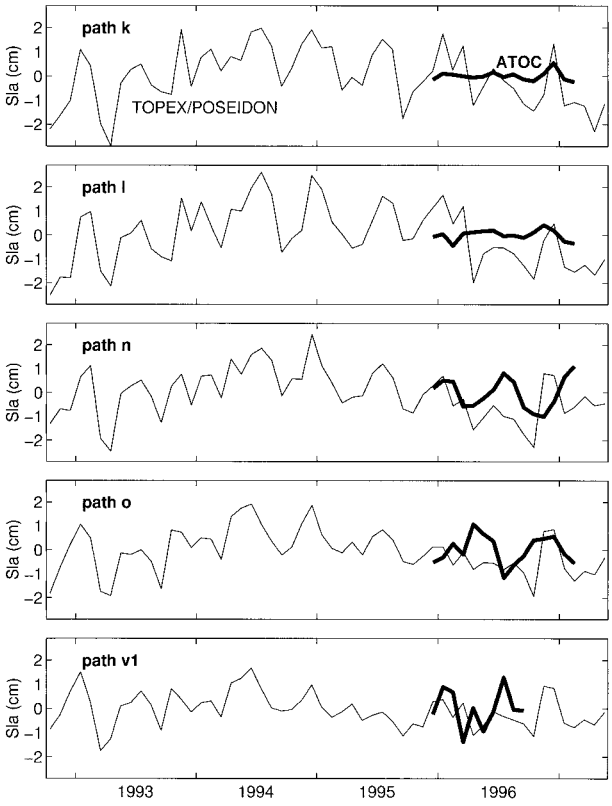

FIG. 8. GCM–ATOC residual, along the five sections shown on Fig. 1, converted to an equivalent sea level anomaly for comparison with the TOPEX/Poseidon data. Annual cycles and trends have been removed.

b. ATOC data

We now turn our attention to the acoustic data. Figure 8 compares the GCM–ATOC residual, converted to an equivalent sea level anomaly, to the range-averaged GCM–TOPEX/Poseidon residual along each acoustic path, after removing trends and annual cycles. The acoustic data is used to estimate the vertical structure of the errors and to test noise model #3 from above, that is, $\mathbf{Q} = 0.047\mathbf{Q}_1$. We model $\mathbf{Q}$ as a diagonal matrix with four parameters, $\alpha_1, \ldots, \alpha_4$, each representing system error variance associated with each of the four vertical EOFs, and with a spatial structure proportional to that of the GCM–TOPEX/Poseidon residual variance (Fig. 6c). Measurement and representation error for the acoustic data are modeled as $\mathbf{R} = \alpha_5\mathbf{I}$.

The cost function (18) is minimized assuming a priori estimates of 0.047 ± 0.047 for $\alpha_1, \ldots, \alpha_4$, that is, the estimate obtained using TOPEX/Poseidon data but allowing for a larger uncertainty in order to test the vertical equipartition hypothesis. The a priori estimate for α_5 is taken to be 0.28 ± 0.28, that is, the variance of the acoustic data with a corresponding uncertainty. A conservative estimate for the prior sample covariance uncertainty is $\mathbf{R}_\epsilon = 0.28\mathbf{I}$ (appendix C). The resulting estimates for system noise variance are $\alpha_1 = 0.15 \pm$

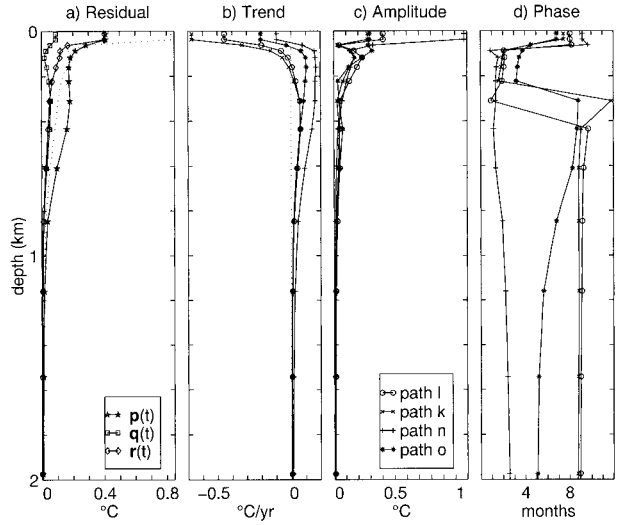

FIG. 9. Vertical structure of the errors along the ATOC sections: (a) standard error (in °C) excluding trend and annual cycle, (b) trend (in °C year^{-1}), (c) annual cycle amplitude (in °C), and (d) annual cycle phase in months. The pentagrams, squares, and diamonds in (a) correspond to estimated GCM, system, and measurement standard errors, respectively. The dotted line is the mean standard uncertainty of the acoustic inversions: it can be compared to that estimated using covariance matching (diamonds) and it provides an approximate measure of statistical significance. Trends and annual cycles are displayed for acoustic sections k, l, n, and o of Fig. 1. Positive trends correspond to warming of the GCM relative to the acoustic data. Annual cycle phase indicates the month of maximum positive anomaly for the GCM relative to the data.

0.04, $\alpha_2 = 0.00 \pm 0.04$, $\alpha_3 = 0.11 \pm 0.04$, and $\alpha_4 = 0.00 \pm 0.04$. These estimates differ from the altimetric estimate of 0.047 ± 0.006, indicating that the vertical equipartition hypothesis is not valid.

A solution that is simultaneously consistent with both TOPEX/Poseidon and ATOC data can also be obtained: $\alpha_1 = 0.04 \pm 0.03$, $\alpha_2 = 0.01 \pm 0.02$, $\alpha_3 = 0.06 \pm 0.03$, and $\alpha_4 = 0.12 \pm 0.02$. This solution differs from that using ATOC data alone in that it predicts less error variance associated with vertical EOF 1 and more with vertical EOF 4, that is, larger model errors above the seasonal thermocline (see Fig. 2). The differences are likely caused by different spatial and temporal extents for the ATOC and TOPEX/Poseidon data and by inaccuracies in the assumed statistical models. All three covariance matching solutions, however, whether from TOPEX/Poseidon data alone, from the ATOC data, or from their combination predict that about 60% of the GCM–TOPEX/Poseidon residual variance is explained by system error.

Figure 9a displays the mean vertical structure of residual errors along the ATOC acoustic sections. The dotted line indicates the mean standard uncertainty of the acoustic inversions (The ATOC Consortium 1998) and can be compared to the covariance matching estimate of $\alpha_5 = 0.31 \pm 0.03$ (diamonds). Also displayed are the estimated GCM and system standard errors, $\mathbf{p}(t)$ and $\mathbf{q}(t)$, respectively. The acoustic data has limited depth resolution, being better suited to the measurement of top-to-bottom averages. Nevertheless, the data indicates significant errors in the GCM variability from about 100-m to 1000-m depth, with a maximum of 0.2°C at 300 m.

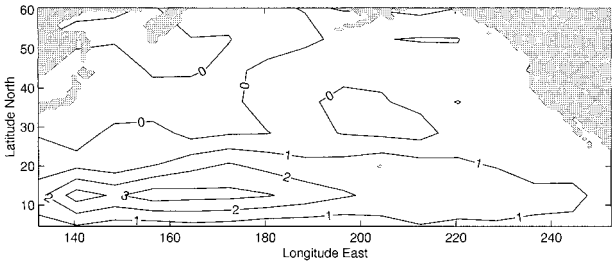

FIG. 10. Trend in the GCM-TOPEX/Poseidon residual. Contour intervals are in cm year^{-1} of sea level anomaly. Positive contours indicate a gradual warming of the GCM relative to TOPEX/Poseidon.

c. *Trend and annual cycle*

Trends and annual cycles of the GCM-data residuals, which were excluded from the previous analysis, are discussed next. In the tropical Pacific, the GCM exhibits a warming trend relative to TOPEX/Poseidon data of up to 3 cm year^{-1} (Fig. 10). The acoustic data indicate that most of the warming occurs between the seasonal and main thermoclines, 50–1000-m depth, with a peak warming of 0.1° to 0.2°C year^{-1}, depending on location (Fig. 9b).

For most of the subtropical gyre, both the GCM and TOPEX/Poseidon exhibit maximum sea level anomaly in September (month 9), but the TOPEX/Poseidon amplitude is about 2 cm larger than that of the GCM (Fig. 11). As a result, the peak GCM–TOPEX/Poseidon residual occurs in March (month 3), six months out of phase with the GCM or TOPEX/Poseidon annual cycle. Excluding the surface layer, where resolution is poor, the acoustic data suggest that the annual cycle error is confined to a depth range shallower than 200 m, the phase-locked range in Fig. 9d, with a peak of 0.3°C at 120-m depth (Fig. 9c).

5. Summary and concluding remarks

The principal contribution of this study is the couching of the GCM error estimation problem in terms of familiar least-squares theory. The so-called ''covariance matching'' approach makes it possible to take advantage of a large number of tools from discrete linear inverse theory in order to study the statistical properties of the errors. The present study demonstrates that covariance matching is both a powerful diagnostic for addressing theoretical questions and an efficient approach for practical applications.

Assuming system and data errors to be uncorrelated

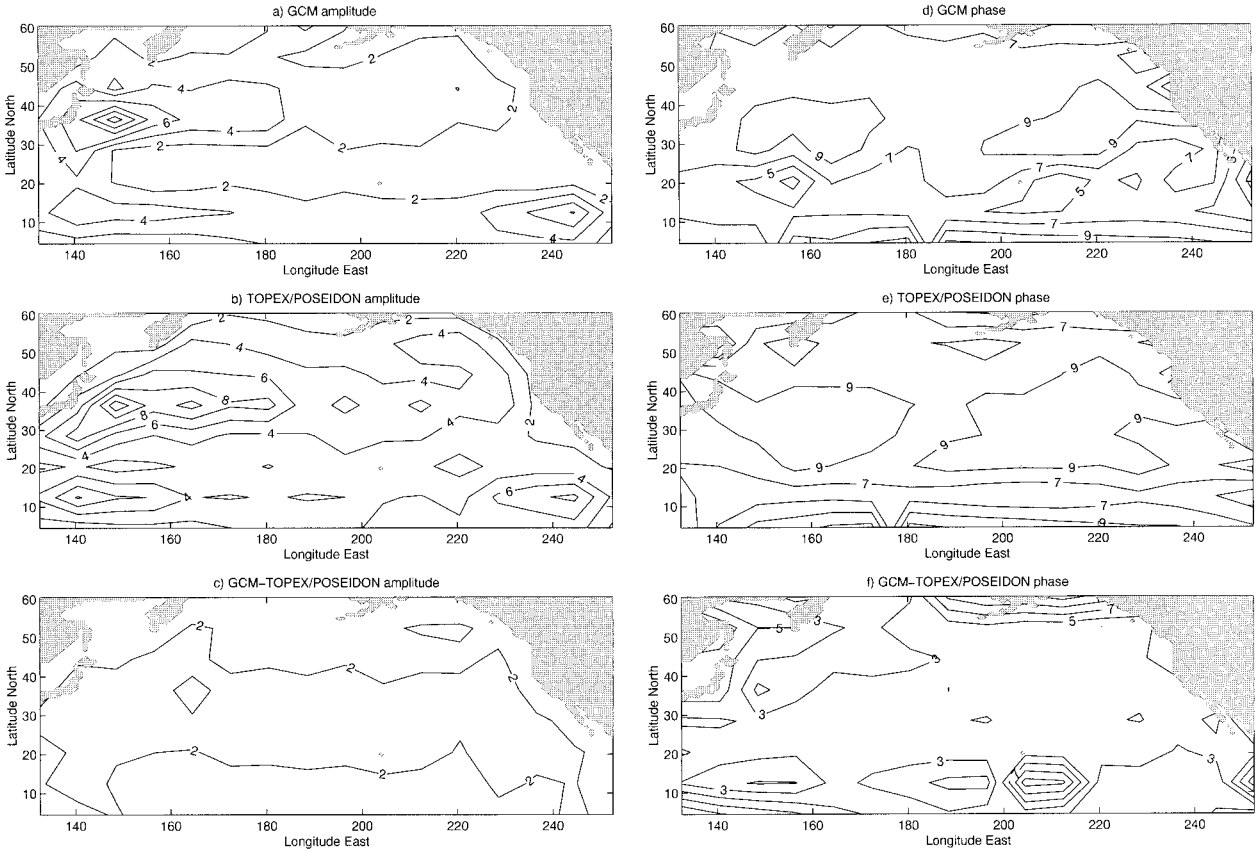

FIG. 11. Annual cycle peak amplitude for (a) GCM output, (b) TOPEX/Poseidon data, and (c) GCM–TOPEX/Poseidon residual. Contour intervals are 2 cm. The corresponding phase is displayed in (d), (e), and (f), respectively, with 2-month contour intervals.

from each other and from the oceanic state, theoretical questions can be addressed in the context of least-squares equations (17)–(18). For a particular GCM and set of measurements, the Green's function matrix, $\mathcal{G}$ in (17), establishes which GCM and data error components are resolvable. Component $\mathbf{Q}_k$ of system error covariance matrix $\mathbf{Q}$ is resolvable provided $\mathbf{H}\mathbf{A}^s\mathbf{Q}_k\mathbf{A}^{s\prime}\mathbf{H}' \neq \mathbf{0}$ for some $s \geq 1$ (section 2b). At least N independent measurements per time step and two covariance matrices (from the set $\mathbf{Y}, \mathbf{D}_1, \mathbf{D}_2, \ldots$) are required to fully resolve an $N \times N$ matrix $\mathbf{Q}$ (appendix B). When matrix $\mathbf{Q}$ is fully observable, then from (8) and (9) the data error covariance matrix $\mathbf{R}$ is also fully observable.

A major obstacle to obtaining statistically significant results is the large uncertainty of sample covariance matrices, $O(2\sigma^4/p)$ where σ^2 is the variance and p is the degrees of freedom (appendix C). The sample uncertainty is represented by $\mathbf{R}_\epsilon$ in (18) and standard least-squares tools can be used to evaluate the statistical significance of the error estimates (sections 2c and 2d). In general, the number of error covariance parameters, α_k in (6)–(7), which can be determined with a reasonable degree of statistical significance is two to three orders of magnitude smaller than the total number of independent data.

We illustrate the approach by applying it to a particular integration of the Marshall et al. (1997a,b) GCM, 56 months of TOPEX/Poseidon sea level anomaly data, and 14 months of acoustic tomography data from the ATOC project. The GCM is forced with observed meteorological conditions at the surface and integrated in a global configuration with 1° horizontal grid spacing and 20 vertical levels. A reduced state linear model that describes internal (baroclinic) error dynamics is constructed for the study area (5°–60°N, 132°–252°E).

Twin experiments, using the reduced state model, suggest that altimetric data are ill-suited to the estimation of covariance matrices for internal GCM errors, but that such estimates can in theory be obtained using the acoustic data (Figs. 4 and 5). These conclusions must however be qualified in the following way. First, the vertical modes used here are EOFs, not dynamical modes, and second, the tests were conducted using linearized GCM dynamics. We do not exclude the possiblity that dynamical modes or fully nonlinear dynamics could enhance the resolution of internal GCM errors from altimetric data.

The GCM exhibits a warming trend relative to TOPEX/Poseidon data of order 1 cm year^{-1} (Fig. 10) corresponding to a peak warming of up to 0.2°C year^{-1} in

the acoustic data at depths ranging from 50 to 200 m (Fig. 9b). This trend measures GCM drift. At the annual cycle, GCM and TOPEX/Poseidon sea level anomaly are in phase, but GCM amplitude is 2 cm smaller (Fig. 11). The acoustic data suggest that the annual cycle error is confined to the top 200 m of ocean (Figs. 9c and 9d). These differences result from errors in the surface boundary conditions and in the dynamics of the GCM.

After removal of trends and annual cycles, the low-frequency/wavenumber (periods >2 months, wavelengths >16°) TOPEX/Poseidon sea level anomaly is order 6 cm^2. The GCM explains about 40% of that variance (Fig. 6). Assuming the error model used to be correct, it is estimated by covariance matching that about 60% of the GCM–TOPEX/Poseidon residual variance is consistent with the reduced state dynamical model (Fig. 7b). This conclusion appears to be relatively robust: it holds for a number of different vertical and horizontal error models, and it is supported by both the altimetric and the acoustic data. The acoustic data measure significant GCM temperature errors in the 100–1000-m depth range with a maximum of 0.3°C rms at 300 m (Fig. 9a). The remaining GCM–TOPEX/Poseidon residual variance is attributed to measurement noise, to barotropic and salinity GCM errors, and to vertical modes of temperature variability that are not represented by the reduced state model.

This and previous studies demonstrate that it *is* possible to obtain simple statistical models for GCM errors that are consistent with the available data. For practical applications, however, the GCM error covariance estimation problem is in general highly underdetermined, much more so than the state estimation problem. In other words there exist a very large number of statistical models that can be made consistent with the available data. Therefore, methods for obtaining quantitative error estimates, powerful though they may be, cannot replace physical insight. But used in the right context, as a tool for guiding the choice of a small number of model parameters, covariance matching can be a useful addition to the repertory of oceanographers seeking to quantify GCM errors or to carry out data assimilation studies.

Acknowledgments. We thank C. Wunsch for providing scientific guidance and commenting on an early draft of the manuscript. We also thank C. Frankignoul and I. Fukumori for insightful discussions. Financial support was provided by SERDP/ARPA as part of the ATOC project (University of California SIO contract PO 10037358) and by NASA Grant NAGW-1048. MC was partially supported by a NASA Global Change Sciences Fellowship.

APPENDIX A

The Fu et al. (1993) Approach

The covariance matching approach of Fu et al. (1993) is derived from (1), (3), and (4):

$$\mathbf{x}_{\mathrm{GCM}}(t) = \mathbf{x}_{\mathrm{ocean}}(t) + \mathbf{p}(t), \tag{A1}$$

$$\mathbf{y}_{\mathrm{ocean}}(t) = \mathbf{H}\mathbf{x}_{\mathrm{ocean}}(t) + \mathbf{r}(t), \tag{A2}$$

$$\mathbf{y}(t) = \mathbf{H}\mathbf{p}(t) - \mathbf{r}(t), \tag{A3}$$

where $\mathbf{p}(t)$ and $\mathbf{r}(t)$ are GCM and data errors, respectively, and $\mathbf{y}(t)$ is the GCM data residual. Multiplying each expression by its transpose and taking expectations yields

$$\mathrm{cov}\mathbf{x}_{\mathrm{GCM}} = \mathrm{cov}\mathbf{x}_{\mathrm{ocean}} + \mathbf{P} + 2\langle \mathbf{x}_{\mathrm{ocean}}\mathbf{p}'\rangle, \tag{A4}$$

$$\mathrm{cov}\mathbf{y}_{\mathrm{ocean}} = \mathbf{H}\,\mathrm{cov}\mathbf{x}_{\mathrm{ocean}}\mathbf{H}' + \mathbf{R}, \tag{A5}$$

$$\mathbf{Y} = \mathbf{H}\mathbf{P}\mathbf{H}' + \mathbf{R}, \tag{A6}$$

where it is assumed that $\mathbf{r}(t)$ is uncorrelated from $\mathbf{x}_{\mathrm{ocean}}(t)$ and from $\mathbf{p}(t)$. Fu et al. (1993) further assume that $\langle \mathbf{x}_{\mathrm{ocean}}\mathbf{p}'\rangle = \mathbf{0}$ and solve for $\mathbf{R}$ and $\mathbf{H}\mathbf{P}\mathbf{H}'$ to obtain equations (32) and (33). Because the approach described in this manuscript does not require $\langle \mathbf{x}_{\mathrm{ocean}}\mathbf{p}'\rangle = \mathbf{0}$, it is possible to evaluate the validity of this assumption:

$$\mathbf{H}\langle \mathbf{x}_{\mathrm{ocean}}\mathbf{p}'\rangle\mathbf{H}' \approx \frac{1}{2}(\mathbf{H}\,\mathrm{cov}\mathbf{x}_{\mathrm{GCM}}\mathbf{H}' - \mathrm{cov}\mathbf{y}_{\mathrm{ocean}} + \tilde{\mathbf{R}} - \mathbf{H}\tilde{\mathbf{P}}\mathbf{H}'). \tag{A7}$$

APPENDIX B

Maximum Number of Resolvable Parameters

The following restrictions apply to the estimation of error covariance matrices $\mathbf{Q}$ and $\mathbf{R}$. First, the maximum number of parameters α_k in (6)–(7) that can be resolved is $M(N + 1)$, where $M \leq N$ is the number of independent observations at a given time step and N is the state dimension [the length of vector $\mathbf{p}(t)$]. Second, the maximum number of parameters α_k describing $\mathbf{Q}$ that can be resolved is $M(N + 1) - M(M + 1)/2$. This restriction applies whether $\mathbf{R}$ is specified a priori or not. Finally, the number of independent parameters α_k, which can be resolved using S data covariance matrices (from the set $\mathbf{Y}$, $\mathbf{D}_1$, $\mathbf{D}_2$, . . .) is at most $SM(M + 1)/2$. Proofs are established for the covariance matching approach by computing the rank of the Green's function matrix in (13) (Chechelnitsky 1999).

Consider an $N \times N$ system error covariance matrix $\mathbf{Q}$ and an $M \times M$ measurement error covariance matrix $\mathbf{R}$. Here $\mathbf{Q}$ and $\mathbf{R}$ are symmetric matrices and are therefore completely described by $N(N + 1)/2$ and by $M(M + 1)/2$ parameters, respectively. To fully resolve $\mathbf{Q}$ and $\mathbf{R}$, the above rules require that N independent observations be available (i.e., $M = N$) and that at least two data covariance matrices (e.g., $\mathbf{Y}$ and $\mathbf{D}_1$) be used.

APPENDIX C

Uncertainty of Sample Covariance

The covariance of a sample covariance (15) is

$$
\begin{aligned}
&(T-q)(T-r)\ \mathrm{cov}[\tilde{\mathbf{Y}}_{(i,j)}(q),\ \tilde{\mathbf{Y}}_{(k,l)}(r)]\\
&\quad=\sum_{s=1}^{T-q}\sum_{u=1}^{T-r}[\mathbf{Y}_{(i,k)}(s-u)\mathbf{Y}_{(j,l)}(s+q-u-r)+\mathbf{Y}_{(i,l)}(s-u-r)\mathbf{Y}_{(j,k)}(s+q-u)]\\
&\quad-\frac{1}{T-r}\sum_{s=1}^{T-q}\sum_{u,v=1}^{T-r}[\mathbf{Y}_{(i,k)}(s-u)\mathbf{Y}_{(j,l)}(s+q-v-r)+\mathbf{Y}_{(i,l)}(s-v-r)\mathbf{Y}_{(j,k)}(s+q-u)]\\
&\quad-\frac{1}{T-q}\sum_{s,t=1}^{T-q}\sum_{u=1}^{T-r}[\mathbf{Y}_{(i,k)}(s-u)\mathbf{Y}_{(j,l)}(t+q-u-r)+\mathbf{Y}_{(i,l)}(s-u-r)\mathbf{Y}_{(j,k)}(t+q-u)]\\
&\quad+\frac{1}{(T-q)(T-r)}\sum_{s,t=1}^{T-q}\sum_{u,v=1}^{T-r}[\mathbf{Y}_{(i,k)}(s-u)\mathbf{Y}_{(j,l)}(t+q-v-r)+\mathbf{Y}_{(i,l)}(s-v-r)\mathbf{Y}_{(j,k)}(t+q-u)]. \qquad \text{(C1)}
\end{aligned}
$$

Here $\mathbf{Y}_{(i,j)}(q)$ denotes the (i, j)th element of the lag-q covariance matrix $\langle[\mathbf{y}(t+q)-\langle\mathbf{y}\rangle][\mathbf{y}(t)-\langle\mathbf{y}\rangle]'\rangle$; $\tilde{\mathbf{Y}}_{(i,j)}(q)$ is the corresponding sample covariance. This formula is a generalization of the univariate expression derived by Anderson [1971, section 8.2, Eq. (64)]. The formula is exact for Gaussian time series. In practice, however, lag covariances in (C1) are replaced by sample estimates, leading to approximate solutions. A useful approximation is

$$
\begin{aligned}
&\mathrm{cov}[\tilde{\mathbf{Y}}_{(i,j)}(0),\ \tilde{\mathbf{Y}}_{(k,l)}(0)]\\
&\quad\approx\frac{1}{p}[\tilde{\mathbf{Y}}_{(i,k)}(0)\tilde{\mathbf{Y}}_{(j,l)}(0)+\tilde{\mathbf{Y}}_{(i,l)}(0)\tilde{\mathbf{Y}}_{(j,k)}(0)], \qquad \text{(C2)}
\end{aligned}
$$

where $p \le T$ is the number of degrees of freedom (roughly, the number of time steps T divided by the e-folding correlation period). From (C2), the variance of a sample covariance is $O[\sigma_1^2\sigma_2^2(1+\rho^2)/p]$, where σ_1^2 and σ_2^2 denote the variances of two random variables and ρ is the correlation coefficient. The probability distribution of a sample covariance is approximately normal for $p > 30$; it is chi-square for $\rho = \pm 1$ (e.g., Mardia et al. 1979). Uncertainty for the lag-s difference covariance matrix (12) can be computed by observing that

$$
\mathbf{D}_s = 2[\mathbf{Y}(0)-\mathbf{Y}(s)]. \qquad \text{(C3)}
$$

(MATLAB-callable software for estimating uncertainty of sample covariance is available via anonymous FTP to gulf.mit.edu, IP Address 18.83.0.149, from directory pub/dimitri/GCMerror.)

REFERENCES

Anderson, B. D. O., and J. B. Moore, 1979: *Optimal Filtering.* Prentice-Hall, 357 pp.

Anderson, T. W., 1971: *The Statistical Analysis of Time Series.* Wiley Series in Probability and Mathematical Statistics, John Wiley & Sons, 704 pp.

ATOC Consortium, 1998: Ocean climate change: Comparison of acoustic tomography, satellite altimetry, and modeling. *Science,* **281,** 1327–1332.

Belanger, P. R., 1974: Estimation of noise covariance matrices for a linear time-varying stochastic process. *Automatica,* **10,** 267–275.

Bennett, A. F., 1992: *Inverse Methods in Physical Oceanography. Cambridge Monogr. on Mechanics and Applied Mathematics,* Cambridge University Press, 346 pp.

Blanchet, I., C. Frankignoul, and M. A. Cane, 1997: A comparison of adaptive Kalman filters for a tropical Pacific Ocean model. *Mon. Wea. Rev.,* **125,** 40–58.

Cane, M. A., A. Kaplan, R. N. Miller, B. Tang, E. C. Hackert, and A. J. Busalacchi, 1996: Mapping tropical Pacific sea level: Data assimilation via a reduced state space Kalman filter. *J. Geophys. Res.,* **101,** 22 599–22 617.

Chechelnitsky, M., 1999: Adaptive error estimation in linearized ocean general circulation models. Ph.D. thesis, Massachusetts Institute of Technology and Woods Hole Oceanographic Institution, 211 pp. [Available from Joint Program in Physical Oceanography, Massachusetts Institute of Technology, Cambridge, MA 02139.]

Cohn, S. E., 1997: An introduction to estimation theory. *J. Meteor. Soc. Japan,* **75,** 257–288.

Daley, R., 1992: The lagged innovation covariance: A performance diagnostic for atmospheric data assimilation. *Mon. Wea. Rev.,* **120,** 178–196.

Dee, D. P., 1995: On-line estimation of error covariance parameters for atmospheric data assimilation. *Mon. Wea. Rev.,* **123,** 1128–1145.

——, and A. M. da Silva, 1998: Data assimilation in the presence of forecast bias. *Quart. J. Roy. Meteor. Soc.,* **124,** 269–295.

——, S. Cohn, A. Dalcher, and M. Ghil, 1985: An efficient algorithm for estimating noise covariances in distributed systems. *IEEE Trans. Autom. Control,* **30,** 1057–1065.

Evensen, G., D. Dee, and J. Schröter, 1998: Parameter estimation in dynamical models. *Ocean Modeling and Parameterizations,* E. P. Chassignet and J. Verron, Eds., Kluwer Academic, 373–398.

Fu, L.-L., and R. D. Smith, 1996: Global ocean circulation from satellite altimetry and high-resolution computer simulation. *Bull. Amer. Meteor. Soc.,* **77,** 2625–2636.

——, I. Fukumori, and R. N. Miller, 1993: Fitting dynamic models to the Geosat sea level observations in the tropical Pacific Ocean. Part II: A linear, wind-driven model. *J. Phys. Oceanogr.,* **23,** 2162–2181.

Fukumori, I., 1999: Assimilation of TOPEX/Poseidon altimeter data into a global ocean circulation model: Are the results any good? *J. Geophys. Res.,* in press.

——, and P. Malanotte-Rizzoli, 1995: An approximate Kalman filter for ocean data assimilation: An example with an idealized Gulf Stream model. *J. Geophys. Res.,* **100** (C4), 6777–6793.

Gajić, Z., and M. T. J. Qureshi, 1995: *Lyapunov Matrix Equation in System Stability and Control.* Academic Press, 255 pp.

Groutage, F. D., R. G. Jacquot, and R. L. Kirlin, 1987: Techniques for adaptive state estimation through the utilization of robust smoothing. *Control and Dynamic Systems,* C. T. Leondes, Ed., Vol. 25, Academic Press, 273–308.

Hayes, S. P., L. J. Mangum, J. Picaut, A. Sumi, and K. Takeuchi,

1991: TOGA–TAO: A moored array for real-time measurements in the tropical Pacific Ocean. *Bull. Amer. Meteor. Soc.,* **72,** 339–347.

Holland, W. R., and P. Malanotte-Rizzoli, 1989: Assimilation of altimeter data into an oceanic circulation model: Space versus time resolution studies. *J. Phys. Oceanogr.,* **19,** 1507–1534.

Isaksson, A., 1988: On system identification in one and two dimensions with signal processing applications. Ph.D. thesis, Linköping University, 249 pp. [Available from Dept. of Electrical Engineering, Linköping University, S-581 83 Linköping, Sweden.]

Malanotte-Rizzoli, P., Ed., 1996: *Modern Approaches to Data Assimilation in Ocean Modeling.* Elsevier Oceanography Series, Vol. 61, Elsevier, 455 pp.

Mardia, K. V., J. T. Kent, and J. M. Bibby, 1979: *Multivariate Analysis.* Academic Press, 521 pp.

Marshall, J., A. Adcroft, C. Hill, L. Perelman, and C. Heisey, 1997a: A finite-volume, incompressible Navier–Stokes model for studies of the ocean on parallel computers. *J. Geophys. Res.,* **102** (C3), 5753–5766.

——, C. Hill, L. Perelman, and A. Adcroft, 1997b: Hydrostatic, quasi-hydrostatic and non-hydrostatic ocean modeling. *J. Geophys. Res.,* **102** (C3), 5733–5752.

Maybeck, P. S., 1979: *Stochastic Models, Estimation, and Control.* Academic Press, 423 pp.

Menemenlis, D., and C. Wunsch, 1997: Linearization of an oceanic circulation model for data assimilation and climate studies. *J. Atmos. Oceanic Technol.,* **14,** 1420–1443.

——, P. W. Fieguth, C. Wunsch, and A. S. Willsky, 1997: Adaptation of a fast optimal interpolation algorithm to the mapping of oceanographic data. *J. Geophys. Res.,* **102** (C5), 10 573–10 584.

Menke, W., 1989: *Geophysical Data Analysis: Discrete Inverse Theory.* International Geophysics Series, Vol. 45, Academic Press, 285 pp.

Moghaddamjoo, R. R., and. R. L. Kirlin, 1993: Robust adaptive Kalman filtering. *Approximate Kalman Filtering,* G. Chen, Ed., World Scientific, 65–85.

Myers, K. A., and B. D. Tapley, 1976: Adaptive sequential estimation with unknown noise statistics. *IEEE Trans. Autom. Control,* **21,** 520–523.

Shellenbarger, J., 1966: Estimation of covariance parameters for an adaptive Kalman filter. *Proc. National Electronics Conf.,* Chicago, IL, National Engineering Consortium, 698–702.

Stammer, D., and C. Wunsch, 1994: Preliminary assessment of the accuracy and precision of TOPEX/Poseidon altimeter data with respect to the large-scale ocean circulation. *J. Geophys. Res.,* **99** (C12), 24 584–24 604.

——, R. Tokmakian, A. Semtner, and C. Wunsch, 1996: How well does a ¼° global circulation model simulate large-scale oceanic observations? *J. Geophys. Res.,* **101** (C11), 25 779–25 811.

Wunsch, C., 1996: *The Ocean Circulation Inverse Problem.* Cambridge University Press, 442 pp.